Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 23 (2013) 53 – 59

Procedia Computer Science

4th International Conference on Computational Systems-Biology and Bioinformatics, CSBio2013

A model of amylopectin biosynthesis based on the activities of key enzymes coupled with the crystallization process

Teeraphan Laomettachit [a,*]

[a]*Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand*

Abstract

Amylopectin is the major glucose polymer in the plant starch granule. The amylopectin molecules have a highly ordered structure of alternating regions between the crystalline and amorphous lamellae, which is distinct from glycogen (the principle form of carbohydrate storage in bacteria and animals). Although several key enzymes involved in the amylopectin biosynthesis have been studied extensively, the underlying mechanism that gives rise to the unique structure of amylopectin remains largely unknown. In the present work, the Monte Carlo simulation technique is used to study the roles of the key enzymes and their relationship in the biosynthesis of amylopectin. We propose a mechanism of the amylopectin biosynthesis based on the activities of the key enzymes coupled with the crystallization process. The proposed mechanism can explain the characteristics of the molecular weight distributions of the molecules and how the amylopectin structure arises.

Keywords: amylopectin; starch biosynthesis; Monte Carlo simulation

1. Introduction

Starch from plants is the major carbohydrate source of calories in the human diet. Plant starch molecules have a highly ordered structure of glucan chains that are packed tightly into the granular form inside the plastids (chloroplasts and amyloplasts). The starch molecules consist of amylose and amylopectin but it is essentially the

* Corresponding author. Tel.: 66-0-2470-7532
E-mail address: teeraphan.lao@mail.kmutt.ac.th

doi:10.1016/j.procs.2013.10.008

amylopectin molecules that contribute to the structure of the granules as amylose synthesis is not required for the formation of the granules[1].

Amylopectin can be distinguished from glycogen, bacterial and animal starch storage, by its distinct pattern of semi-crystallization. The pattern results from the packing of the amylopectin molecules in the starch granules, which is organized on several scales (see [2] for a review). At the lowest level of the amylopectin structure, glucosyl units are linked by α(1-4) linkages into linear glucan chains. These linear chains are joined at branching points by α(1-6) linkages. The pattern of branching is not random but is regularly ordered into clusters. This pattern of the amylopectin structure is believed to enable the storage of larger amounts of glucose molecules, compared to the structure of bacterial and animal glycogen that possesses more random branches, which would lead to steric hindrance that prevents further growth of the molecules.

The most accepted view of the amylopectin structure is the cluster model[2,3] (Figure 1). In this model, nonrandom branching positions of α(1-6) linkages are resided in the amorphous region occupying ~3 nm of the ~9-nm repeating distance of the cluster unit. The other ~6-nm part of the unit is composed of α(1-4) linear chains that are packed tightly by forming double helices of adjacent chains, allowing the parallel alignment structure to form the crystalline lamellae. The clustered branch-structure of the molecules is believed to allow higher organization of the starch although the mechanisms are less understood.

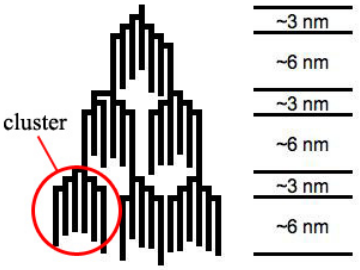

Fig. 1. Cluster model of the amylopectin molecule

The enzymes participating in the biosynthesis of amylopectin have been studied extensively[3,4]. Three major families of the enzymes are starch synthases (SSs), branching enzymes (BEs), and debranching enzymes (DBEs). SSs elongate the glucan chains by adding a glucosyl unit to the non-reducing end of the growing chains by forming the α(1-4)-glucosyl bond. BEs cleave the existing α(1-4) linkage and transfer the cut segment to another or the same glucan chain to form the α(1-6)-glucosyl bond, creating a new branch of the molecule. DBE was initially believed not to involve in the amylopectin biosynthesis pathway. However, *dbe* mutants from many plants show dramatic structure change of the amylopectin molecules to a more randomly branched structure called phytoglycogen, suggesting an important role of DBE in the amylopectin biosynthesis. Several models have proposed the roles of DBE that involve either directly[3] or indirectly[5] in the biosynthesis of amylopectin.

Although the enzymatic activities of the amylopectin biosynthesis have been characterized, the underlying mechanism that gives rise to the distinct structure of amylopectin remains largely unknown. This is because the enzymes, in fact, act cooperatively and the functions of the individual enzymes cannot lead to the understanding of the amylopectin biosynthesis process. Rather, a system-level study of the relationship between enzymes' activities and the crystallization of the molecules should be investigated.

In the present paper, a modeling framework for studying the roles of the key enzymes and their relationship in the biosynthesis of amylopectin is developed. We then propose a model of interplays between the activities of SS and BE, and the crystallization process of the molecules. Our simple model is able to explain the characteristics of the molecular weight distributions of the molecules as measured in the experiment[6].

2. Models and Results

2.1. Model 1 (SS+BE)

In Model 1, we investigate the roles of SS and BE in the biosynthesis process of amylopectin. Our model keeps track of the lengths of all chains in the molecule and how the chain lengths are modified by the enzymes. The chain length can be referred to as the degree of polymerization (DP), which is the number of glucosyl units in the chain. The simulation outcome of the model is the chain DP profile of the molecule, which can be plotted as the molecular weight distribution.

SS and BE in our model represent the collective isoforms of the starch synthases and branching enzymes, respectively. SS adds a glucosyl unit to an existing chain, therefore, extending the chain length by one. BE cuts existing chains that have the length above a threshold. The threshold is set at 12 DP according to the minimal DP preference of BE[7]. The DP units of the cut segment are randomly determined, given that both the residual and the cut segments must be at least 6 DP. Once BE catalyzes the reaction, the chain length of the donor chain is reduced by the randomly determined DP and a new chain with the same determined DP is added to the list. It should be noted that the position of the newly created chain in the molecule is not of interest in the current study. In other words, the effect of steric interference between branches is ignored.

We start our simulation with a pre-existing chain of 6 DP. The model allows only one reaction to occur in each simulation step. We use Monte Carlo simulation to determine which enzymatic reaction occurs on which chain. Chains with DP<12 are associated only with the elongation reaction catalyzed by SS whereas chains with DP≥12 are associated with both the elongation (catalyzed by SS) and the branching (catalyzed by BE) reactions. Therefore, the first six reaction steps of the simulation are always the extension of the pre-existing chain catalyzed by SS. At the seventh reaction step, the pre-existing chain can either get elongated by SS or get branched by BE.

Here we demonstrate how the reaction is chosen. In Table 1, a chain length state of amylopectin is shown. The molecule consists of three chains with the lengths of 15, 8 and, 7. The parent ID indicates that the chains with ID=2 and ID=3 were both created from the chain with ID=1 by the branching reaction. The chain with ID=1 is the pre-existing chain.

Table 1. An example of the chain length state of the amylopectin molecule

Chain ID	Parent ID	Chain Length
1	–	15
2	1	8
3	1	7

Table 2. All possible reactions associated with the amylopectin molecule state in Table 1

Reaction	Substrate chain ID	Reaction type	Probability
1	1	Elongation	$p_1 = p_{SS}$
2	1	Branching	$p_2 = p_{BE}$
3	2	Elongation	$p_3 = p_{SS}$
4	3	Elongation	$p_4 = p_{SS}$
			$p_{sum} = p_{1+} p_{2+} p_{3+} p_4$

For the current simulation step, there are four possible events that can take place (Table 2). The chain with ID=1 has the length ≥12 DP, therefore, it can either get catalyzed by SS (with probability p_1/p_{sum}) or by BE (with probability p_2/p_{sum}), where $p_{sum}=p_1+p_2+p_3+p_4$. The chains with ID=2 and ID=3 have their lengths <12 DP, therefore, they can only get catalyzed by SS with probability p_3/p_{sum} and p_4/p_{sum}, respectively. p_{SS} and p_{BE} are model parameters and we set p_{SS}:p_{BE}=10:1.

The reaction is chosen randomly according to the calculated probability above. For example, if the reaction 1 is chosen the length of the chain with ID=1 becomes 16. If the reaction 2 is chosen, the program will randomly choose the DP of the cut segment between 6 and 9 (note again that, the minimum length of any chains is 6 DP). For instance if the chosen length is 8, the length of the chain with ID=1 becomes 15–8=7 and a new chain with DP=8 is added to the table. The new chain is assigned ID=4 and parent ID=1. The probability of each possible reaction is then recalculated for the next simulation step. The molecular state of amylopectin is kept modified by this rule until the program is terminated.

We simulate the model until the molecule has 10,000 chains and repeat the simulation 30 times. We plot the molecular weight distribution of the molecules in Figure 2 (blue dots), where N is the degree of polymerization (DP) and $P(N)$ is the number of chains with DP=N. The distribution can be compared directly to the experiment[6] (Figure 2, lines) to validate our model.

In the experiment[6], the amylopectin molecules were treated with debranching enzymes, which cleaved all the branching points of the molecules but left individual chain lengths unperturbed. Debranched amylopectin chains were then separated by their molecular weights (i.e., lengths) using fluorophore-assisted capillary electrophoresis. The molecular weight distributions of amylopectin from various plants are shown in Figure 2.

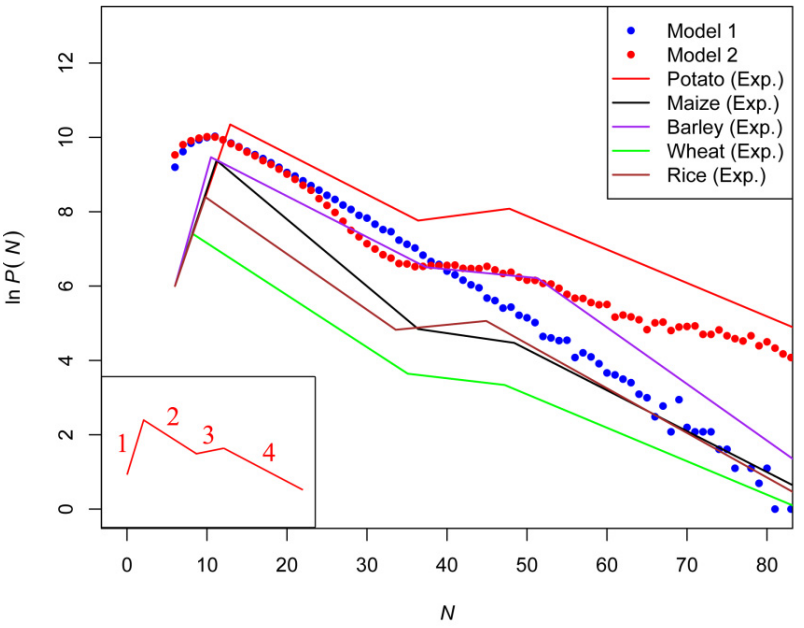

Fig. 2. Molecular weight distributions of starch from models and plants

The molecular weight distributions of amylopectin from various plants reveal four characteristic regions (Figure 2, inset). As discussed in[6], the values of the negative slope in the regions 2 and 4 reflect relative rates of chain growth and chain growth inhibition. In our model, the growth of the chains is due to the SS activity and the inhibition of the chain growth is due the activity of BE. The regions of the chain lengths with the relatively higher activity of BE compared to the activity of SS (i.e., low growth rate) have more negative values of the slopes. The positive or small negative slopes observed in the regions 1 and 3 infer different mechanisms other than simple chain growth and chain growth inhibition. Castro *et al*.[6] discussed that the region 1 corresponds to the formation of new branches while the region 3 corresponds to the formation of the crystalline structure of the molecule.

The activities of SS and BE in our Model 1 can only produce the regions 1 and 2 (Figure 2). In fact, the molecular weight distribution of the Model 1 molecules resembles that of glycogen[6]. Since amylopectin is distinct from glycogen by its crystallization structure, we investigate the crystallization process of the molecules in the following Model 2 in order to explain the appearance of the regions 3 and 4.

2.2. Model 2 (SS+BE+crystallization)

In Model 2, we investigate the hypothesis that the amylopectin structure arises from the interplays between the activities of the enzymes and the crystallization of the molecules. In this model, we assign a cluster ID to each chain. New chains created are assigned the same cluster ID as of their parents, indicating that they will be in the same cluster. We assume that once the number of chains with the same cluster ID exceeds a threshold, the chains can crystallize to form a cluster (the threshold is assumed at 30 chains in the model). Once the crystallization has taken place, the chains in the cluster become water insoluble and inaccessible by SS and BE, and therefore cannot be further modified by the enzymes.

However, it is essential for the molecule to have some chains remain active so that the molecule can grow further. We assume that a few chains are partially crystallized and are, therefore, able to elongate beyond the current cluster to seed the next cluster. The mechanism of how these chains are selected is not known but it was suggested in[6] that chains with their lengths not fitting in a cluster size should be able to elongate further. Following the idea, we randomly select 1-3 chains from those with the longest lengths in the cluster to extend further. A new cluster ID is then assigned to these chains.

How do the regions 3 and 4 of amylopectin appear? We explain this by the following mechanism: when the cluster crystallization takes place, chains with short lengths (e.g., chains in regions 1 and 2) are packed and become water insoluble and inaccessible by the enzymes. Chains with longer chain lengths (those seen in region 3) are partially crystallized, leaving a few glucosyl residuals at the end of the chains remain water soluble and active. We assume that BE can access only to the water-soluble end residuals of these chains. Since the end residual lengths are short (below the minimal threshold of BE; DP=12), these residuals are protected from the BE activity. Therefore, chains with the lengths seen in the region 3 can be catalyzed only by SS, making the growth rate of this region relatively high. This explains why the region 3 shows positive slopes (in potato and rice) or small negative slopes (in maize, barley, and wheat). Until the residuals further elongate to have the residual end DP≥12 (chain lengths seen in region 4), BE can resume its activity making the chains in this region grow slower and have relatively larger negative slope values.

The molecular weight distribution simulated from Model 2 is plotted in Figure 2 (red dots) along with the distributions of amylopectin from other plants. The distribution of our hypothetical starch molecule from Model 2 shows the unique four characteristics and has the slope values comparable to other plants (Table 3).

3. Discussion and Conclusions

The amylopectin molecules are distinct from glycogen in their crystalline structure. As discussed in[3], there are two possible scenarios for the crystallization process. First, glucan chains are synthesized and modified by the enzymes and the crystallization process happens subsequently at once. Second, the synthesis and modifications of the glucan chains occur concurrently with the crystallization process.

Table 3. Characteristic slopes calculated from our models and from experiment[6]

Plant source	Region 1	Region 2	Region 3	Region 4
Model 1	0.32	–0.13	n/a	n/a
Model 2	0.19	–0.16	–0.01	–0.08
Potato	0.63	–0.11	0.028	–0.09
Maize	0.65	–0.18	–0.031	–0.11
Barley	0.77	–0.11	–0.021	–0.15
Wheat	0.64	–0.14	–0.025	–0.09
Rice	0.63	–0.15	0.021	–0.12

Our model supports the latter hypothesis. The actions of SS and BE prepare the glucan chains for the crystallization process. In turn, the crystallization prevents the glucan chains from enzyme accessibility by forming the crystalline cluster. Only some glucan chains in the cluster can escape from the crystallization to reinitiate the next cluster.

We develop a simple model of the amylopectin biosynthesis based on the activities of starch synthases and branching enzymes coupled with the crystallization of the molecules. Our simulation shows that the four region characteristics of the molecular weight distributions can be explained by the proposed mechanism in the model. The regions 1 and 2 observed in both Model 1 and Model 2 are the results of the combination actions of SS and BE. The regions 3 and 4 (observed in Model 2) are the results of the crystallization process. In our model, the crystallization process packs the chains into the crystalline cluster and prevents most of the chains from further modifications by the enzymes. The chains that are able to escape the crystallization process (presumably the longest chains in the cluster) protrude part of the chains out of the crystalline region. This part of the protruding chain is subject to the minimal DP requirement (DP≥12) by BE. Our model suggests that the regions 3 and 4 arise from the combination actions of SS and BE similar to what is observed in the regions 1 and 2, but the actions occur on the residual parts of the longer chains that protrude from the previous cluster.

Our simple model can capture the essential characteristics of the molecular weight distributions of the molecules and shed light on the mechanism underlying the amylopectin biosynthesis process. In the future work, the details of multiple SS and BE isoforms specific to different plants will be studied in order to fit the model to each experimental data. In addition, the role of the debranching enzymes in the amylopectin biosynthesis process remains to be investigated.

Acknowledgements

This research is supported by the “KMUTT Research Fund” from King Mongkut’s University of Technology Thonburi, Thailand.

References

1. Ball S, Guan H-P, James M, Myers A, Keeling P, Mouille G, Buléon A, Colonna P, Preiss J. From glycogen to amylopectin: a model for the biogenesis of the plant starch granule, *Cell*. 1996;**86**:49-352.

2. Pérez S, Bertoft E. The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review, *Starch - Stärke*. 2010;**62**:389-420.

3. Myers AM, Morell MK, James MG, Ball SG. Recent progress toward understanding biosynthesis of the amylopectin crystal, *Plant physiology*. 2000;**122**:989-997.

4. Nakamura Y. Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue, *Plant and cell physiology*. 2002;**43**:718-725.

5. Streb S, Delatte T, Umhang M, Eicke S, Schorderet M, Reinhardt D, Zeeman SC. Starch granule biosynthesis in Arabidopsis is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase, *The Plant Cell Online*. 2008;**20**:3448-3466.
6. Castro JV, Dumas C, Chiou H, Fitzgerald MA, Gilbert RG. Mechanistic information from analysis of molecular weight distributions of starch, *Biomacromolecules*, 2005;**6**:2248-2259.
7. Nakamura Y, Utsumi Y, Sawada T, Aihara S, Utsumi C, Yoshida M, Kitamura S. Characterization of the reactions of starch branching enzymes from rice endosperm, *Plant and cell physiology*, 2010;**51**:776-794.